ELSEVIER
Discrete Applied Mathematics 50 (1994) 77–88
DISCRETE APPLIED MATHEMATICS

On the complexity of the dual method for maximum balanced flows

Uwe Zimmermann

Abteilung für Mathematische Optimierung, Institut für Angewandte Mathematik, Technische Universität Carolo-Wilhelmina, Pockelsstraße 14, D-3300 Braunschweig, Germany

Received 16 September 1991; revised 1 April 1992, 17 July 1992

Abstract

In an earlier paper we develop a quite general dual method and apply it to balanced submodular flow problems with flow values in modules. Here, we analyze that method for the particular case of balanced flows with rational or integral flow values in more detail. While, for integral flows, the general problem turns out to be NP-hard, the method is strongly polynomial for rational as well as for integral flows when applied to the motivating reliability problem given by Minoux. In that case, a maximum balanced flow is determined in $O(m \cdot M(m, n))$, where $M(m, n)$ is the complexity of some maxflow procedure for a network with n vertices and m arcs.

1. Introduction

When a line connecting two relay stations in a telephone network fails due to some technical problem then all telephone calls routed via that line are lost. In order to assure some reliability level, the number of lost calls should be less than a given proportion of the total number of calls routed in the network. In 1976, Minoux [6] modeled such a situation as a maximum flow problem with additional capacity constraints for all arcs e:

$$x(e) \leqslant \sigma \cdot v(x),$$

where $x(e)$ is the flow value in arc e in the network, $v(x)$ the total flow from source to sink, and σ, $0 < \sigma < 1$, is the reliability level. For given σ, the maximum number of calls that can be routed through the network is the maximum total flow from source to sink. A solution to the maximum flow problem also defines a possible routing of the calls.

Minoux [6] derives a pseudopolynomial algorithm for solving that reliability problem. In [8], we develop a dual method for solving *maximum balanced flow problems*, i.e. for maximizing the total flow value subject to additional upper bounds of the form

SSDI 0166-218X(92)00150-U

$$x(e) \leqslant \alpha(e) \cdot v(x) + \beta(e), \tag{1}$$

on all arcs e of the network. The method is described in a quite general setting for submodular flows with flow values in modules. A theoretical variant of the dual method is equivalent to a maxflow procedure working with parametric data. In [8], that method is shown to be strongly polynomial for rational flows. The approach is of theoretical interest only and is not recommended for practical implementation. For integral flows, the approach fails. In the same paper, we conjecture the problem to be NP-hard for integral flows. Lower bounds of the form

$$\alpha'(e) \cdot v(x) - \beta'(e) \leqslant x(e) \tag{2}$$

can be added without changing these results, as shown in [1].

In Section 2, we define the maximum balanced flow problem and give a short proof of its NP-hardness in the presence of lower and upper bounds. The idea for the proof is drawn from [1]. Without lower bounds of the form (2), the complexity of the problem remains open, although still conjectured to be NP-hard. In Section 3, we shortly describe the dual method in a form suitable for complexity analysis. Here, the method generates a strictly decreasing sequence of upper bounds, where each bound is calculated from a minimum arc or a minimum cut generated by some maxflow algorithm. For rational flows pseudopolynomial time bounds are derived in Section 4. In the particular case of the above reliability problem, one of these bounds implies an $O(m \cdot M(m, n))$ time bound on the dual method, where $M(m, n)$ denotes the complexity of the underlying maxflow procedure for a graph with n vertices and m arcs.

For integral flows, pseudopolynomial time bounds are trivial. In Section 5, we assume that no bounds of the form (2) occur and discuss some results on the effective calculation of upper bounds in the dual method. Again, an exact calculation is seemingly difficult but is not known to be NP-hard. On the other hand, in the particular case of the above reliability problem, bound calculation turns out to be easy, and we succeed in deriving the same time bound as for rational flows, i.e. $O(m \cdot M(m, n))$.

In order to compare the theoretical results with computational experience we briefly review some experience from [1] in Section 6.

2. Balanced flows

We consider flows with rational or integral values on the arcs of directed graphs (digraphs). Let $\mathbb{N}$ denote the set of all natural numbers, let $\mathbb{Z}$ ($\mathbb{Z}_+$) denote the set of all (nonnegative) integral numbers, and let $\mathbb{Q}$ ($\mathbb{Q}_+$) denote the set of all (nonnegative) rational numbers.

Let $G = (V, E)$ denote a digraph with vertex set V and arc set E. For $S \subseteq V$, let $\bar{S} := V \backslash S$ and let $\delta(S) := \{ij \in E \mid i \in S, j \in \bar{S}\}$. All arcs have lower and upper capacities, denoted by $l(e)$ and $u(e)$, $e \in E$. For $x : E \to \mathbb{Q}$ and $A \subseteq E$, let $x(A) := \sum_{e \in A} x(e)$.

A *rational* (*integral*) *flow* in G is a function $x: E \to \mathbb{Q}(\mathbb{Z})$ satisfying all *capacity constraints*, i.e. $l \leqslant x \leqslant u$, where $l \leqslant u: E \to \mathbb{Q}(\mathbb{Z})$, as well as the *flow conservation constraints*, i.e. $x(\delta(\bar{v})) - x(\delta(v)) = 0$, for all $v \in V$. In particular for integral flows, all capacities are assumed to be integral, since rational capacities could be rounded to integral ones without changing the set of all integral flows.

Let $\hat{e} \in E$ denote some fixed arc in G and let α, α', β, $\beta': E\backslash\{\hat{e}\} \to [0, 1)$, where $[0, 1) := \{r \in \mathbb{Q}_+ \mid r < 1\}$. Then a flow is called *balanced* if it satisfies the *balancing constraints*, i.e. for all $e \in E\backslash\{\hat{e}\}$:

$$\alpha'(e) \cdot x(\hat{e}) - \beta'(e) \leqslant x(e) \leqslant \alpha(e) \cdot x(\hat{e}) + \beta(e). \tag{3}$$

For integral flows, we will also round these capacities, but since they depend on the flow value $x(\hat{e})$, the rounded capacities depend on $x(\hat{e})$, too. We remark that all results in the following sections easily generalize to balancing problems in which lower and/or upper balancing constraints are only defined for subsets of $E\backslash\hat{e}$.

Balancing assures that flow values lie within some interval around certain proportions of the flow on the special arc $\hat{e}$. Minoux [6] introduces the problem in the particular case where $l = 0$, $\alpha' = \beta' = \beta = 0$, and where α is constant. He describes a pseudopolynomial method for solving the corresponding *maximum balanced flow problem*:

$$z_* := \max\{x(\hat{e}) \mid x \text{ balanced flow}\}. \tag{4}$$

In [8], we develop a dual method for the more general balanced submodular flow problem. Although described only for constant lower bounds, the method can easily be extended to lower bounds of the above type (cf. [1]). For rational flows, the method solves the maximum balanced flow problem in a finite number of steps and allows a strongly polynomial variant. For integral flows even with constant lower bounds the problem is conjectured to be NP-hard in [8].

In fact, for integral flows with general lower bounds of the above type, the maximum balanced flow problem is NP-hard. For arbitrary rational coefficients in the balancing constraints, Ahlers [1] proposed a reduction of the *good simultaneous approximation problem*. Since we allow only rational coefficients from [0, 1), a refinement of that reduction is necessary. In the following, we describe the refined reduction.

Simultaneous diophantine approximation (SDA).

Input: $\alpha \in \mathbb{Q}^n$; $0 < \varepsilon \in \mathbb{Q}$; $N \in \mathbb{N}$.

Question: Do there exist $y \in \mathbb{N}$ with $1 \leqslant y \leqslant N$ and $x \in \mathbb{Z}^n$ such that

$$|\alpha_i \cdot y - x_i| \leqslant \varepsilon$$

for all i, $1 \leqslant i \leqslant n$?

Lagarias [4] has shown that SDA is NP-complete.

Theorem 2.1. *Deciding whether there exists a balanced integral flow is* NP-*complete.*

Proof. It suffices to reduce SDA to some balanced flow problem. Obviously, solving SDA is equivalent to solving the set of linear inequalities

$$1 \leqslant y \leqslant N, \quad \text{and,}\ \alpha_i \cdot y - \varepsilon \leqslant x_i \leqslant \alpha_i \cdot y + \varepsilon \quad \text{for all } i.$$

For $\varepsilon \geqslant 1/2$, there are trivially solvable by $y := 1$, $x_i := \lceil \alpha_i - \varepsilon \rceil$. Thus, we may assume w.l.o.g. $\varepsilon < 1/2$. For $\alpha_i \geqslant 1$, we observe equivalence of the corresponding ith inequality to

$$(\alpha_i - \lfloor \alpha_i \rfloor) \cdot y - \varepsilon \leqslant x_i - \lfloor \alpha_i \rfloor \cdot y \leqslant (\alpha_i - \lfloor \alpha_i \rfloor) \cdot y + \varepsilon.$$

Thus, w.l.o.g. we assume $0 \leqslant \alpha_i < 1$ for all i. Owing to these assumptions, any solution x, y satisfies

$$0 \leqslant \lceil \alpha_i \cdot y - \varepsilon \rceil \leqslant x_i \leqslant \lfloor \alpha_i \cdot y + \varepsilon \rfloor \leqslant \lfloor y + 1/2 \rfloor = y,$$

implying $0 \leqslant x_i \leqslant y \leqslant N$.

We define the corresponding balanced flow problem in the digraph $G(V, E)$ with vertex set $V := \{s, t\}$ and with arc set $E := \{\hat{e}, e_1, e_2, \ldots, e_{n+1}\}$, where all arcs e_i, $i = 1, \ldots, n+1$ lead from s to t, while $\hat{e}$ leads from t to s. Capacity and balancing constraints are given by

$$\alpha_i \cdot x(\hat{e}) - \varepsilon \leqslant x(e_i) \leqslant \alpha_i \cdot x(\hat{e}) + \varepsilon, \quad 0 \leqslant x(e_i) \leqslant N,$$

for all $i = 1, \ldots, n$, and by

$$-\, n \cdot N \leqslant x(e_{n+1}) \leqslant N, \quad 1 \leqslant x(\hat{e}) \leqslant N.$$

Identifying $x(\hat{e}) \equiv y$ and $x(e_i) \equiv x_i$, one easily verifies that an integral balanced flow subject to these constraints exists if and only if the corresponding linear inequalities admit an integral solution. Nakayama [7] observed that it is impossible to define a suitable balancing constraint on e_{n+1}. If we insist on balancing constraints on all arcs different from $\hat{e}$, we have to replace the arc e_{n+1} by one arc a_0 from s to t and n arcs $a_1, \ldots, a_n$ from t to s with suitable capacities and balancing constraints given by

$$0 \leqslant x(a_i) \leqslant \frac{N}{N+1} \cdot x(\hat{e}) + \frac{N}{N+1}, \quad 0 \leqslant x(a_i) \leqslant N,$$

for all $i = 0, \ldots, n$. $\square$

By inspection of the proof, we observe that feasibility for integral balanced flows is NP-complete even if $\alpha = \alpha'$, $\beta = \beta'$ provided that we do not insist on balancing constraints on all arcs different from $\hat{e}$. On the other hand, the conjectured NP-completeness in the special case $\alpha' = \beta' = 0$ still remains to be investigated.

3. The dual method

In this section we describe the dual method from [8] specialized for solving balanced flow problems as defined in the previous section. For fixed parameter z, $l(\hat{e}) \leqslant z \leqslant u(\hat{e})$, the existence of some balanced flow with $x(\hat{e}) = z$ is equivalent to the existence of some flow in G subject to the parametric capacity constraints

$$l(e,z) \leqslant x(e) \leqslant u(e,z) \tag{5}$$

for all $e \in E$ where the parametric lower and upper bounds are defined by

$$l(\hat{e},z) := z, \qquad u(\hat{e},z) := z,$$

and, for all $e \in E \setminus \{\hat{e}\}$, by

$$l(e,z) := \max\{\alpha'(e)\cdot z - \beta'(e), l(e)\},$$

$$u(e,z) := \min\{\alpha(e)\cdot z + \beta(e), u(e)\}.$$

As for the nonparametric bounds, for $x(\cdot,z): E \to \mathbb{Q}$ and $A \subseteq E$, let $x(A,z) := \sum_{e\in A} x(e,z)$.

In the integral case, we use the same notation for the suitably rounded bounds. Since, in this case, we assumed that all capacity constraints are integral, we have

$$l(e,z) := \max\{\lceil \alpha'(e)\cdot z - \beta'(e)\rceil, l(e)\},$$

$$u(e,z) := \min\{\lfloor \alpha(e)\cdot z + \beta(e)\rfloor, u(e)\}.$$

Then, in any case, feasibility is characterized in terms of Hoffman's well-known feasibility criterium for network flows (cf. Lawler [5]), i.e. for all $S \subseteq V$,

$$0 \leqslant f(S,z) := u(\delta(S),z) - l(\delta(\bar{S}),z), \tag{6}$$

and for all $e \in E$,

$$0 \leqslant f(e,z) := u(e,z) - l(e,z). \tag{7}$$

Let $Z(S)$, for $S \subseteq V$, $[Z(e)$, for $e \in E]$ denote the sets of all parameters z for which the corresponding equation (6) [equation (7)] is satisfied. In particular, $Z(\hat{e}) = [l(\hat{e}), u(\hat{e})]$. Then, the set of all parameters Z_{feas} which admit a balanced flow is given by

$$Z_{\text{feas}} = \bigcap_{e\in E} Z(e) \cap \bigcap_{S\subseteq V} Z(S). \tag{8}$$

In the rational case, but not in the integral case, all these sets are convex. We denote the maximal elements in these sets by $\bar{z}(e)$, for $e \in E$, and by $\bar{z}(S)$, for $S \subseteq V$. By definition, the maximal element of an empty set is $-\infty$. Then, obviously,

$$z_* := \max\{x(\hat{e}) \mid x \text{ balanced flow}\} \leqslant \min\left\{\min_{e\in E} \bar{z}(e), \min_{S\subseteq V} \bar{z}(S)\right\}. \tag{9}$$

Because of convexity, for rational flows the weak duality inequality (9) turns out to be tight (cf. [8, 1]). For integral flows there is a duality gap, in general. A simple counterexample is given in [8]. Nevertheless, z_* can iteratively be found via calculation of slightly modified bounds.

Dual method.

Step 0. $\bar{z} := u(\hat{e})$; $z_{\text{old}} := u(\hat{e})$.

Step 1. $f(\hat{e},\bar{z}) := \min\{f(e,\bar{z}) \mid e \in E\}$. If $f(\hat{e},\bar{z}) < 0$ then

$$\bar{z} := \max\{z \mid f(\hat{e},z) \geqslant 0,\ l(\hat{e}) \leqslant z \leqslant \bar{z}\}.$$

Step 2. $f(\hat{S}, \bar{z}) := \min\{f(S, \bar{z}) \mid S \subseteq V\}$. If $f(\hat{S}, \bar{z}) < 0$ then

$$\bar{z} := \max\{z \mid f(\hat{S}, z) \geqslant 0,\ l(\hat{e}) \leqslant z \leqslant \bar{z}\}.$$

Step 3. If $\bar{z} = z_{\text{old}}$ then stop (optimal); if $\bar{z} = -\infty$ then stop (infeasible); $z_{\text{old}} := \bar{z}$; go to step 1.

Let $n := |V|$ and $m := |E|$. The condition in Step 1 can easily be tested in $O(m)$, simultaneously for all arcs. The condition in Step 2 can be tested in polynomial time, say in $O(M(m, n))$, simultaneously for all subsets, using some standard network flow procedure (cf. [2]) which either constructs a feasible flow for the current parameter value or returns a minimum cut $\hat{S}$. The complexity of the method is thus determined by the number of the necessary bound calculations and by the complexity of these calculations. Here, the discussion is different for rational and integral flows and will be given in the following two sections separately.

4. Rational balanced flows

In this section we discuss the complexity of the dual method in the particular case of rational flows. For any arc $e \in E \backslash \hat{e}$, the function

$$u(e, z) = \begin{cases} \alpha(e) \cdot z + \beta(e), & z < z_u(e), \\ u(e), & z \geqslant z_u(e), \end{cases}$$

where the *cornerpoint* $z_u(e)$ is defined by

$$z_u(e) := \begin{cases} (u(e) - \beta(e))/\alpha(e), & \alpha(e) > 0, \\ -\infty, & \alpha(e) = 0,\ \beta(e) \geqslant u(e), \\ \infty, & \alpha(e) = 0,\ \beta(e) < u(e), \end{cases}$$

is piecewise linear and concave with at most one cornerpoint in the interval $Z(\hat{e})$. Analogously, the function $-l(e, z)$ defines a cornerpoint $z_l(e)$ and is also piecewise linear and concave with at most one cornerpoint in $Z(\hat{e})$. Obviously, $u(\hat{e}, z)$ and $-l(\hat{e}, z)$ are linear functions. Therefore, $f(e, z)$, for all arcs $e \in E$, and $f(S, z)$, for all subsets $S \subseteq V$, are piecewise linear concave functions with at most two and $m - 1$ cornerpoints, respectively, in $Z(\hat{e})$.

Let $z_1 < z_2 < \cdots < z_k$ for some $k \leqslant 2m - 2$ denote the ordered sequence of all cornerpoints in the open core of $Z(\hat{e})$. Then, all above functions are linear on the induced $k + 1$ subintervals.

Calculation of new bounds in the dual method is easy. Let $g(z)$ denote the function considered, i.e. $g(z) \equiv f(\hat{e}, z)$ or $g(z) \equiv f(\hat{S}, z)$. Evaluation of g can be done in $O(m)$. Now, $g(\bar{z}) < 0$, and one has to calculate the maximum of $\{z \mid g(z) \geqslant 0,\ l(\hat{e}) \leqslant z \leqslant \bar{z}\}$, i.e. the largest zero in that set. By successively evaluating g at cornerpoints $z_i < \bar{z}$ of g, scanning from the largest to the smallest, one either defects infeasibility, i.e. $g(z) < 0$ for $z < \bar{z}$, or locates the largest zero by finding the first two successive cornerpoints $z_i < z_j$ of g with $g(z_i) \geqslant 0 > g(z_j)$. During the algorithm, in total, any cornerpoint has to be scanned at most once before location of a zero, therefore, the complexity of the bound calculation is $O(m)$ per bound plus $O(m^2)$, in total.

Step 1 of the dual method can be canceled after a suitable reduction of the interval $Z(\hat{e})$ to

$$Z(\hat{e}) := [l(\hat{e}), u(\hat{e})] := \bigcap_{e \in E} Z(e),$$

which may be performed in a preprocessing step in $O(m)$.

In order to bound the number of iterations, one observes that in any subinterval the number of different zeroes for linear functions with rational data is pseudopolynomially bounded. A tighter but in general still only pseudopolynomial bound, is given in the following theorem.

Theorem 4.1. *For some rational number* $r \in (0, 1)$, *let* $\alpha(e), \alpha'(e) \in r \cdot Z_+$ *for all* $e \in E \setminus \hat{e}$, *and let* κ *be the cardinality of the set* $\bigcup_{e \in E \setminus \hat{e}} \{\alpha(e)\} \cup \bigcup_{e \in E \setminus \hat{e}} \{\alpha'(e)\} \cup \{0\}$ *of real numbers. Then the dual method stops after at most* $O(\min\{2^n, \lfloor m/r \rfloor, (1 + (2m-2)/(\kappa - 1))^{\kappa - 1}\})$ *iterations. If* $\alpha' \equiv 0$, *it stops after at most* $O(\min\{2^n, \lfloor 1/r \rfloor, (1 + (m-1)/(\kappa - 1))^{\kappa - 1}\})$ *iterations.*

Proof. The number of subsets of V is 2^n, and, due to the convexity of all $Z(S)$, no set can define a new bound more than once unless infeasibility is detected. If the problem is feasible, then, again due to convexity (cf. [8, 1]), the final bound is optimal, which implies the first part of the bound.

Let $\bar{z}$ denote the current bound with $z_{i-1} < \bar{z} \leqslant z_i$ for some i, $1 \leqslant i \leqslant k + 1$, where $z_0 := l(\hat{e})$, $z_{k+1} := u(\hat{e})$. All functions $f(S, z)$ are linear on that fixed subinterval of adjacent cornerpoints, say $f(S, z) = a_i(S) \cdot z + b_i(S)$. If a new bound $\bar{z}_1$ is calculated from $f(S, \bar{z}_1) = 0$, then $f(S, \bar{z}) < 0$, which implies $a_i(S) < 0$.

Let two successive new bounds $\bar{z}_1 > \bar{z}_2$ be calculated, say, from S and T, and let $z_{j-1} < \bar{z}_1 \leqslant z_j$ for some $j \leqslant i$. By concavity,

$$0 = f(S, \bar{z}_1) \leqslant f(S, \bar{z}) + a_i(S)(\bar{z}_1 - \bar{z}),$$

$$f(S, \bar{z}) \leqslant f(T, \bar{z}) \leqslant f(T, \bar{z}_1) + a_j(T)(\bar{z} - \bar{z}_1) < a_j(T)(\bar{z} - \bar{z}_1),$$

which implies $a_i(S) < a_j(T) < 0$.

Now, $a_i(S)$ is the sum of $\alpha(e)$, $e \in \delta(S)$, minus the sum of some $\alpha'(e)$, $e \in \delta(\bar{S})$, plus λ with $\lambda \in \{0, 1, -1\}$, depending on the location of the arc $\hat{e}$ relative to S.

Therefore, $a_i(S) \in [-m, m]$. Since all proportionality factors $\alpha(e)$, $\alpha'(e)$ are integral multiples of some rational number $r \in (0, 1)$, no more than $3 \cdot \lfloor m/r \rfloor$ different negative values of $a_i(S)$ are possible, implying the second part of the bound. If $\alpha' \equiv 0$, $a_i(S) \geqslant -1$, and negativity requires $\lambda = -1$. Thus, in this case, no more than $\lfloor 1/r \rfloor + 1$ different negative values are possible.

On the other hand, there are $\kappa - 1$ different possible nonzero values for the proportionality factors, which appear on $m_1, m_2, \ldots, m_{\kappa-1}$ corresponding arcs. Obviously, the number of different combinations of these values is bounded by $\prod_{j=1}^{\kappa-1} (m_j + 1)$. Since $\sum_{j=1}^{\kappa-1} m_j$ is bounded by the sum of the corresponding arcs, i.e. by $2m - 2$, we observe $\sum_{j=1}^{\kappa-1} (m_j + 1) \leqslant 2m - 2 + \kappa - 1$. By the inequality of the geometric and arithmetic mean, $(1 + (2m - 2)/(\kappa - 1))^{\kappa - 1}$ is a valid bound on the product, and this implies the last part of the bound. If $\alpha' \equiv 0$, $\sum_{j=1}^{\kappa-1} m_j$ is bounded by $m - 1$, and negativity requires $\lambda = -1$. Thus, the product is bounded by $(1 + (m - 1)/(\kappa - 1))^{\kappa - 1}$, which implies the sharpened last bound. □

In particular, the above theorem provides a strongly polynomial time bound on the dual method for Minoux's [3] original motivating problem.

Corollary 4.2. *For some rational number* $r \in (0, 1)$, *let* $\alpha(e)$, $\alpha'(e) \in \{r, 0\}$. *Then the complexity of the dual method is* $O(m \cdot M(m, n))$. *If* $\alpha' \equiv 0$, *the complexity reduces to* $O(\min\{\lfloor 1/r \rfloor, m\} \cdot M(m, n))$.

Proof. As $\kappa = 2$, the above theorem assures termination after $O(m)$ iterations. Bound calculation runs in $O(m)$ per bound plus $O(m^2)$ in total. Thus, the total time bound, including preprocessing, i.e. recalculation of $Z(\hat{e})$ and ordering the cornerpoints, is $O(m \cdot M(m, n))$. In the case $\alpha' \equiv 0$, the improved bound directly follows from the corresponding improved bound of the theorem. □

Furthermore, the method is strongly polynomial whenever $1/r$, where r is the common divisor of the proportionality factors α and α', is polynomially bounded in m and/or n.

We remark that in a forthcoming paper [9] polynomiality of the dual method for solving rational flow problems is provided in general.

5. Integral balanced flows

Here, the optimum value of the corresponding rational balanced flow problem can be used as initial upper bound in the dual method. Since the new bound in any nonfinal iteration is strictly smaller than the old bound, the method stops after at most $u(\hat{e}) - l(\hat{e}) + 1$ iterations. In a straightforward manner the new bound can be calculated by successively testing nonnegativity of the function $g(z)$ considered for $z = \bar{z}, \bar{z} - 1, \ldots$. Since $g(z)$, where $g(z) \equiv f(\hat{e}, z)$ or $g(z) \equiv f(\hat{S}, z)$, can be evaluated in $O(m)$, the dual method is obviously pseudopolynomial. Here, rounding of rational numbers is counted as one elementary operation. As the general problem is NP-hard, there is not much hope for substantial improvement of this bound. In the special case $\alpha' = \beta' = 0$ the complexity of the problem is not known; we also conjecture that it remains NP-hard. In this section, we will only discuss this special case.

Due to the assumption $\alpha' = \beta' = 0$, $l(e, z) = l(e)$, for all $e \in E \backslash \hat{e}$. The functions $u(\hat{e}, z)$ and $l(\hat{e}, z)$ are linear, as in the rational case. For all $e \in E \backslash \hat{e}$, the function

$$u(e, z) = \begin{cases} \lfloor \alpha(e) \cdot z + \beta(e) \rfloor, & z < z_u(e), \\ u(e), & z \geqslant z_u(e), \end{cases}$$

where the *cornerpoint* $z_u(e)$ is defined by

$$z_u(e) := \begin{cases} \lceil (u(e) - \beta(e))/\alpha(e) \rceil, & \alpha(e) > 0, \\ -\infty, & \alpha(e) = 0,\ u(e) = 0, \\ \infty, & \alpha(e) = 0,\ u(e) \geqslant 1, \end{cases}$$

is an isotone, nonnegative function with at most one cornerpoint in the interval $Z(\hat{e})$. Therefore, $f(e, z) = u(e, z) - l(e)$, for all $e \in E \backslash \hat{e}$, is an isotone function and the corresponding set $Z(e)$ is an interval of integers of the form $[\gamma(e), \infty)$. Similar to the rational case, Step 1 of the dual method can be canceled after reduction of the interval $Z(\hat{e})$. Here, only its lower bound may change, i.e. $l(\hat{e}) := \max\{l(\hat{e}), \max\{\gamma(e) \mid e \in E \backslash \hat{e}\}\}$, which can be calculated in $O(m)$.

For all $S \subseteq V$ with $e \notin \delta(\bar{S})$, the function $f(S, z)$ is isotone (cf. definition in inequality (6)). Therefore, if $f(\hat{S}, \bar{z}) < 0$, for some $\hat{S} \subseteq V$ with $\hat{e} \notin \delta(\bar{\hat{S}})$, the problem is infeasible.

Otherwise, for all $S \subseteq V$ with $\hat{e} \in \delta(\bar{S})$, $f(S, z) = -z + \Phi(S, z)$, where

$$\Phi(S, z) := u(\delta(S), z) - l(\delta(\bar{S}) \backslash \hat{e})$$

is an isotone function.

Lemma 5.1. *If $f(\hat{S}, \bar{z}) < 0$ for some $\hat{S} \subseteq V$ with $\hat{e} \in \delta(\bar{\hat{S}})$, then $z \leqslant \Phi(\hat{S}, \bar{z}) < \bar{z}$ for all $z < \bar{z}$ with $f(\hat{S}, z) \geqslant 0$.*

Proof. By assumption, $\Phi(\hat{S}, \bar{z}) < \bar{z}$. If $f(\hat{S}, z) < 0$ for all $z \in \mathbb{Z}$ with $z < \bar{z}$, the first claimed inequality is trivial. Otherwise, let $f(\hat{S}, z) \geqslant 0$ for some $z \in \mathbb{Z}$ with $z < \bar{z}$, i.e. $z \leqslant \Phi(\hat{S}, z)$. By isotonicity, $\Phi(\hat{S}, z) \leqslant \Phi(\hat{S}, \bar{z})$. □

We remark that the use of Φ for the iterative calculation of new bounds in Step 2 of the dual method turns out to be very effective in actual implementations of the method (cf. Section 6).

Let $z_1 < z_2 < \cdots < z_k$, for some $k \leqslant m - 1$, denote the ordered sequence of all cornerpoints in the open core of $Z(\hat{e})$. Then, on any of the induced $k + 1$ subintervals, Φ is a function of the form

$$\phi(z) = c + \sum_{e \in A} \lfloor \alpha(e) \cdot z + \beta(e) \rfloor, \tag{10}$$

with $A \subseteq \delta(\hat{S})$ and $c \in Z$. For such functions, we have to determine

$$\hat{z} := \max\{z \mid z \leqslant \phi(z),\ z \leqslant \bar{z}\}. \tag{11}$$

In general, the exact calculation of $\hat{z}$ is possible in pseudopolynomial time, but not known to be NP-hard. Obviously, calculation of $\hat{z}$ is a special case of the calculation of a maximum balanced integral flow. Therefore, NP-hardness of the former problem would imply the conjectured NP-hardness of the latter problem.

Theorem 5.2. *Let* $a := \sum_{e \in A} \alpha(e)$, *let* $b := \sum_{e \in A} \beta(e)$, *and let* $\hat{z} := \max\{z \mid z \leqslant \phi(z), z \leqslant \bar{z}\}$. *Then*

$$\hat{z} \leqslant \begin{cases} \min[\phi(\bar{z}), \bar{z}], & a \geqslant 1, \\ \min[\phi(\bar{z}), \bar{z}, z_{\max}], & a < 1, \end{cases}$$

where

$$z_{\max} := \left\lfloor \frac{c+b}{1-a} \right\rfloor = c + \left\lfloor \frac{ac+b}{1-a} \right\rfloor$$

If, for some $r, s \in [0, 1)$, $\alpha(e) = r$, $\beta(e) = s$ *for all* $e \in A$, *then* $z_{\max}$ *can be decreased to*

$$z_{\max} := c + |A| \left\lfloor \frac{rc+s}{1-a} \right\rfloor$$

and the inequality becomes tight.

Proof. At first, we consider validity of the upper bounds. By definition of $\hat{z}$, $\hat{z} \leqslant \bar{z}$ and $\hat{z} \leqslant \phi(\hat{z})$. By monotonicity, $\hat{z} \leqslant \phi(\hat{z}) \leqslant \phi(\bar{z})$. It remains to show $\hat{z} \leqslant z_{\max}$, provided that $a < 1$. Here,

$$\phi(\hat{z}) - c = \sum_{e \in A} \lfloor \alpha(e) \cdot \hat{z} + \beta(e) \rfloor \leqslant a \cdot \hat{z} + b \leqslant a \cdot \phi(\hat{z}) + b$$

implies the second inequality in $\hat{z} \leqslant \phi(\hat{z}) \leqslant \lfloor (c+b)/(1-a) \rfloor = z_{\max}$.

Secondly, for constant $\alpha \equiv r$ and $\beta \equiv s$, we consider validity of the obviously decreased bound. Let $v := |A|$. Then $\lfloor r\hat{z} + s \rfloor \leqslant r\phi(\hat{z}) + s = r \cdot (c + v\lfloor r\hat{z} + s \rfloor) + s$ implying

$$\lfloor r\hat{z} + s \rfloor \leqslant \left\lfloor \frac{rc+s}{1-rv} \right\rfloor,$$

which leads to $\hat{z} \leqslant \phi(\hat{z}) = c + v\lfloor r\hat{z} + s \rfloor \leqslant c + v\lfloor (rc+s)/(1-rv) \rfloor$.

Because of the assumptions made, we can also show that the bounds are tight. If the minimum is attained for $\bar{z}$, then $\bar{z} \leqslant \phi(\bar{z})$, which implies $\bar{z} \leqslant \hat{z}$. If $rv = a \geqslant 1$, we show $\phi(\bar{z}) \leqslant \phi(\phi(\bar{z}))$:

$$\begin{aligned} \phi(\phi(\bar{z})) - c &= v\lfloor r\phi(\bar{z}) + s \rfloor \\ &= v\lfloor rc + s + rv\lfloor (r\bar{z} + s \rfloor \rfloor \\ &\geqslant v\lfloor r\bar{z} + s \rfloor \\ &= \phi(\bar{z}) - c. \end{aligned}$$

Therefore, if the minimum is attained for $\phi(\bar{z})$, then $\phi(\bar{z}) \leqslant \hat{z}$. If $rv = a < 1$, let γ denote the minimum. If suffices to prove the claim $\phi(\gamma) \geqslant \gamma$, provided that $\gamma < \bar{z}$. Then, $\gamma = c + v\kappa$, where $\kappa = \min\{\lfloor r\bar{z} + s \rfloor, \lfloor (rc+s)/(1-a) \rfloor\}$, which implies $rc + s \geqslant (1-a)\kappa$. Therefore,

$$\begin{aligned} \phi(\gamma) - c &= v\lfloor r\gamma + s \rfloor = v\lfloor rc + s + rv\kappa \rfloor \\ &\geqslant v\lfloor (1-a)\kappa + a\kappa \rfloor = v\kappa = \gamma - c, \end{aligned}$$

which proves the claim. $\square$

In order to show that bounds derived from the representation $\phi(z)$ of $\Phi(\hat{S}, z)$ on some subinterval can in fact be used as new upper bounds in the dual method, let

$$\phi_i(z) = c_i + \sum_{e \in A_i} \lfloor \alpha(e) \cdot z + \beta(e) \rfloor$$

denote the representation of some $\Phi(\hat{S}, z)$ on the ith subinterval of cornerpoints, say for $z_{i-1} < z \leqslant z_i$, where $1 \leqslant i \leqslant k+1$, and where $z_0 := l(\hat{e})$, $z_{k+1} := u(\hat{e})$. Then,

$$\Phi(\hat{S}, z) = \min\{\phi_i(z) \mid 1 \leqslant i \leqslant k+1\}$$

for all z in the open core of $Z(\hat{e})$. Thus, the above bounds for functions of the form $\phi(z)$ yield corresponding upper bounds from $\Phi(\hat{S}, z)$, even when the subinterval changes. We observe, that with decreasing index i, c_i decreases whereas A_i increases.

In particular, for Minoux's [6] original motivating problem, a new upper bound in the dual method can be calculated in $O(m)$ times the number of touched subintervals. In general, however, we do not know a better than the trivial pseudopolynomial time bound.

For the number of bounds appearing in the dual method, the situation is quite similar. Again, in general, no good complexity bounds are known. Fortunately, in the case of Minoux's [6] original problem, the method turns out to be strongly polynomial.

Theorem 5.3. *If, for some $r, s \in [0, 1)$, $\alpha(e) = r$, $\beta(e) = s$ for all $e \in A$, then the dual method terminates in $O(m \cdot M(m, n))$.*

Proof. Assume that two successive new bounds $\bar{z} > \hat{z}_1 > \hat{z}_2$ are calculated from $\Phi(\hat{S}_1, z)$ and $\Phi(\hat{S}_2, z)$. Let $z_{j-1} < \bar{z} \leqslant z_j$ and $z_{i-1} < \hat{z}_1 \leqslant z_i$ denote the corresponding subintervals, and let $\phi_v(z) = c_v + |A_v| \lfloor rz + s \rfloor$, $\psi_v(z) = d_v + |B_v| \lfloor rz + s \rfloor$, for $i \leqslant v \leqslant j$, denote the corresponding representations of $\Phi(\hat{S}_1, z)$ and $\Phi(\hat{S}_2, z)$. Then

$$\Phi(\hat{S}_2, \bar{z}) \geqslant \Phi(\hat{S}_1, \bar{z}) \geqslant \Phi(\hat{S}_1, \hat{z}_1) \geqslant \hat{z}_1 > \Phi(\hat{S}_2, \hat{z}_1).$$

Due to the identical rounded parts of the functions, such a crossing of function values can only appear when $|B_i| > |A_j|$. Assume on the contrary, $|B_i| \leqslant |A_j|$. Then, due to the above observed monotonicity in the representation of such functions, we have

$$|A_i| \geqslant |A_v| \geqslant |A_j| \geqslant |B_i| \geqslant |B_v| \geqslant |B_j|$$

for all v, $i \leqslant v \leqslant j$, which, due to $\Phi(\hat{S}_2, \bar{z}) \geqslant \Phi(\hat{S}_1, \bar{z})$, implies $\Phi(\hat{S}_2, \hat{z}_1) \geqslant \Phi(\hat{S}_1, \hat{z}_1)$, a contradiction.

Therefore, the factor in front of the rounded part strictly increases with each new bound, i.e. after at most m iterations the dual method stops. Bound calculation within a subinterval can be done in $O(m)$, bound calculation leading into another subinterval can, over all iterations, add no more than $O(m^2)$. Thus, in total, bound calculations can be performed in $O(m^2)$, and the complexity of the method is determined by the $O(m)$ calls to the maxflow procedure generating the necessary minimum cuts in $O(M(m, n))$. □

6. Some remarks on computational experience

In a study at the Technical University Braunschweig, Ahlers [1] reports on computational experience for several classes of rational and integral balanced flow problems. Problems considered are generated with random data and/or using NETGEN [3], with arbitrary lower proportionality terms as well as with $\alpha' \equiv \beta' \equiv 0$. Implementations of the dual method were run in different computing environments, in particular at the AMDAHL mainframe of the computing center and on personal computing equipment of the Abteilung für Mathematische Optimierung.

All problems were solved in 3–10 iterations.

For rational problems as well as for integral problems with trivial lower bounds $\alpha' \equiv \beta' \equiv 0$, the computational burden of an iteration is proportional to the effort for solving the occurring min cut problem. The effectiveness of the method in the integral case seems to be due to the effective iterative bound calculation as discussed in Section 5.

On the contrary, for integral problems with arbitrary lower bounds α', β', when bound calculation more or less consists in testing feasibility of every successive integer, bound calculation clearly dominates the computational effort spent in an iteration. In nclusion, the running time per iteration increases considerably when compared to rational problems. For example, for problems from [3] with about 400 nodes and with about 2400 arcs, an iteration may slow down by a factor of 100 on the AMDAHL mainframe.

References

[1] C. Ahlers, Balanzierte Flüsse, Diploma Thesis, Abteilung für Mathematische Optimierung, Technische Universität Braunschweig (1990).
[2] R.K. Ahuja, Th.L. Magnanti and J.B. Orlin, Network Flows: Theory, Algorithms and Applications (Prentice-Hall, Englewood Cliffs, NJ, 1993).
[3] D. Klingman, A. Napier and J. Stutz, Netgen: a program for generating large scale capacitated assignment, transportation, and minimum cost flow network problems, Management Sci. 20 (1974) 814–821.
[4] J.C. Lagarias, The computational complexity of simultaneous diophantine approximation problems, SIAM J. Comput. 14 (1985) 196–209.
[5] E.L. Lawler, Combinatorial Optimization, Networks and Matroids (Holt, Rinehart & Winston, San Francisco, CA, 1976).
[6] M. Minoux, Flots équilibré et flots avec sécurité, Bull. Direction Etudes Recherches Sér. C. Math. Informat. (1976) 5–16.
[7] A. Nakayama, Comments on Zimmermann's paper: On the complexity of the dual method for maximum balanced flows, Discussion paper, Department of Administration and Social Sciences, Fukushima University (1992).
[8] U. Zimmermann, Duality for balancing submodular flows, Discrete Appl. Math. 15 (1986) 365–376.
[9] U. Zimmermann, Sharing and balancing, in preparation.